Title: PERFORMANCE, POWER, AND ENERGY OF IN-SITU AND POST-PROCESSING VISUALIZATION

Author(s): Adhinarayanan, Vignesh

Intended for: Web

Issued: 2015-10-05

Performance, Power, and Energy of In-situ and Post-Processing Visualization

A Case Study in Climate Simulation

Vignesh Adhinarayanan

Ph.D. (CS) Student, Virginia Tech

GRA, CCS-7 Division, LANL

Virginia Tech Invent the Future

Los Alamos NATIONAL LABORATORY EST. 1943

INTRODUCTION

"Supercomputers are power constrained"

- Power budget for Los Alamos county = 66 MW
- Power budget for Trinity supercomputer alone = 15 MW

- Exceeding power budget ➔ Brownouts in Los Alamos
  - *Installing and starting ASCI White believed to play a part in the rolling California brownouts in 2001*

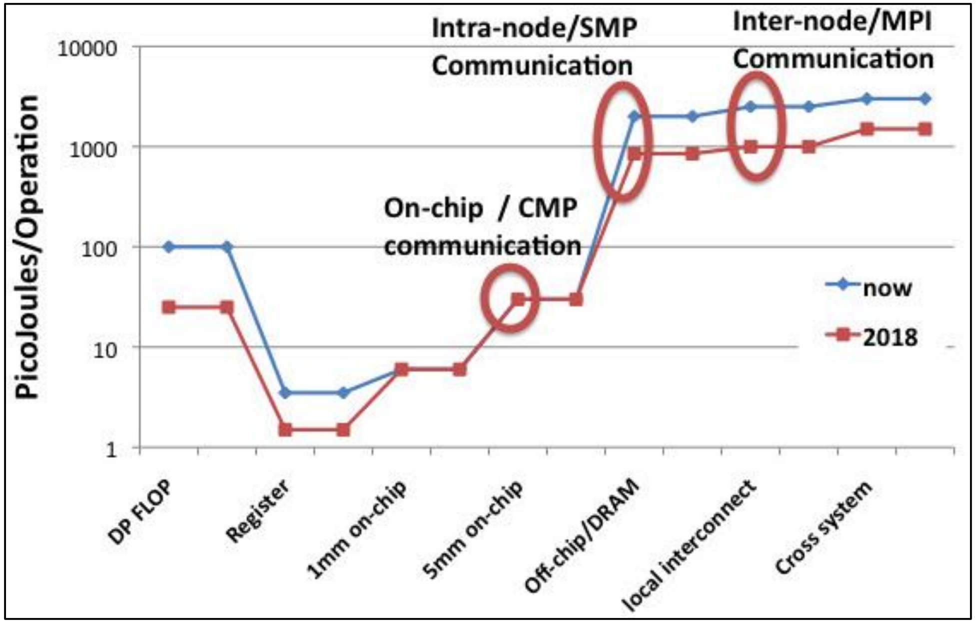
INTRODUCTION

"Supercomputers are energy constrained"

- 1 MW power consumption → 1 million dollars per year
  - Operating cost of supercomputers comparable to the acquisition cost
    - The gap expected to narrow down in the future

WHAT DOES IT HAVE TO DO WITH THE VISUALIZATION TEAM?

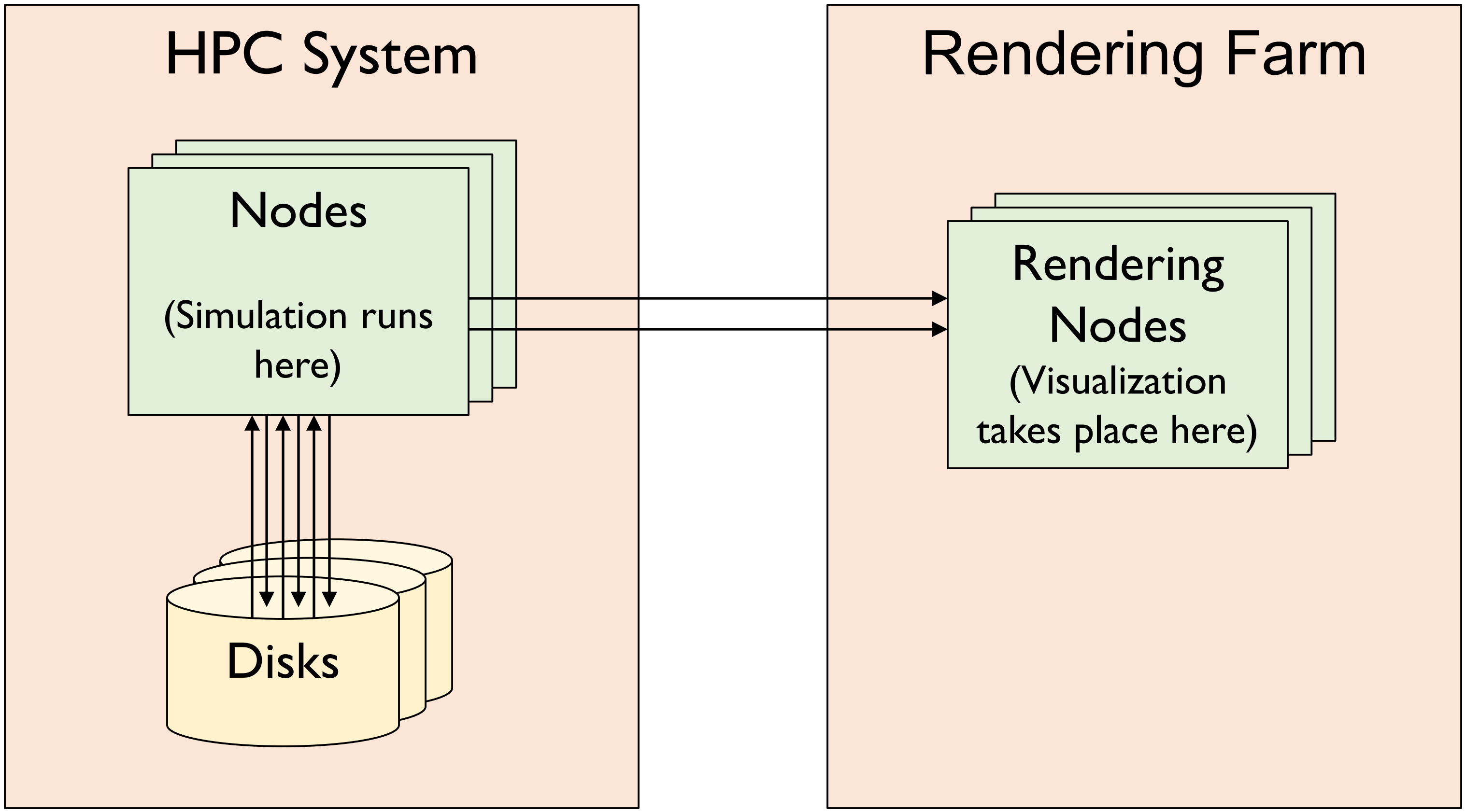
POWER/ENERGY CHALLENGE



- Energy consumed for moving a bit increases as we move down the memory hierarchy
  - Off-chip transfers costs nearly 100 times as much energy as on-chip transfers

Image source: J.Shalf et al., "Exascale Computing Technology Challenges", VECPAR 2010

TRADITIONAL "POST-PROCESSING" VISUALIZATION



Solution: "In-situ" visualization

- In-situ visualization: perform visualization alongside the simulation
  - That is, create an image representation of data at end of each iteration directly instead of writing raw data to the disk
  - The image (reduced size representation) may be written to the disk
  - May involve additional sampling strategies (spatial, temporal, etc.)

OUTLINE

- Introduction
- Objective
- Methodology
- Results
- Conclusion and Future Work

GOAL

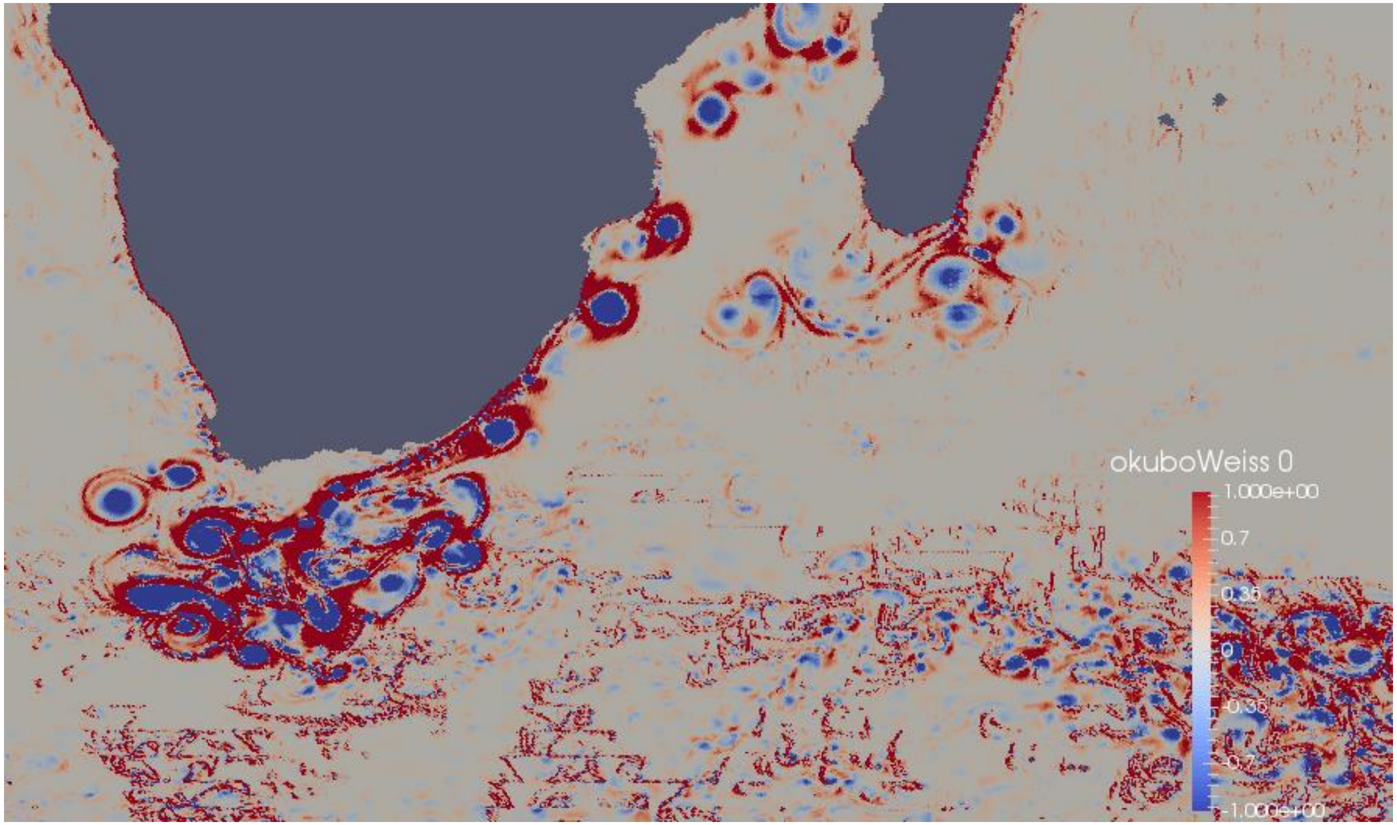
"Study the performance, power, and energy trade-offs among traditional post-processing, modern post-processing, and in-situ visualization pipelines"

- Detailed sub-component level power measurements within a node to gain detailed insights
  - i.e., measure power consumption of CPU, memory, and disk
- Measurements at scale to understand problems unique to big supercomputers

Outline

- Introduction
- Objective
- Methodology
- Results
- Conclusion and Future Work

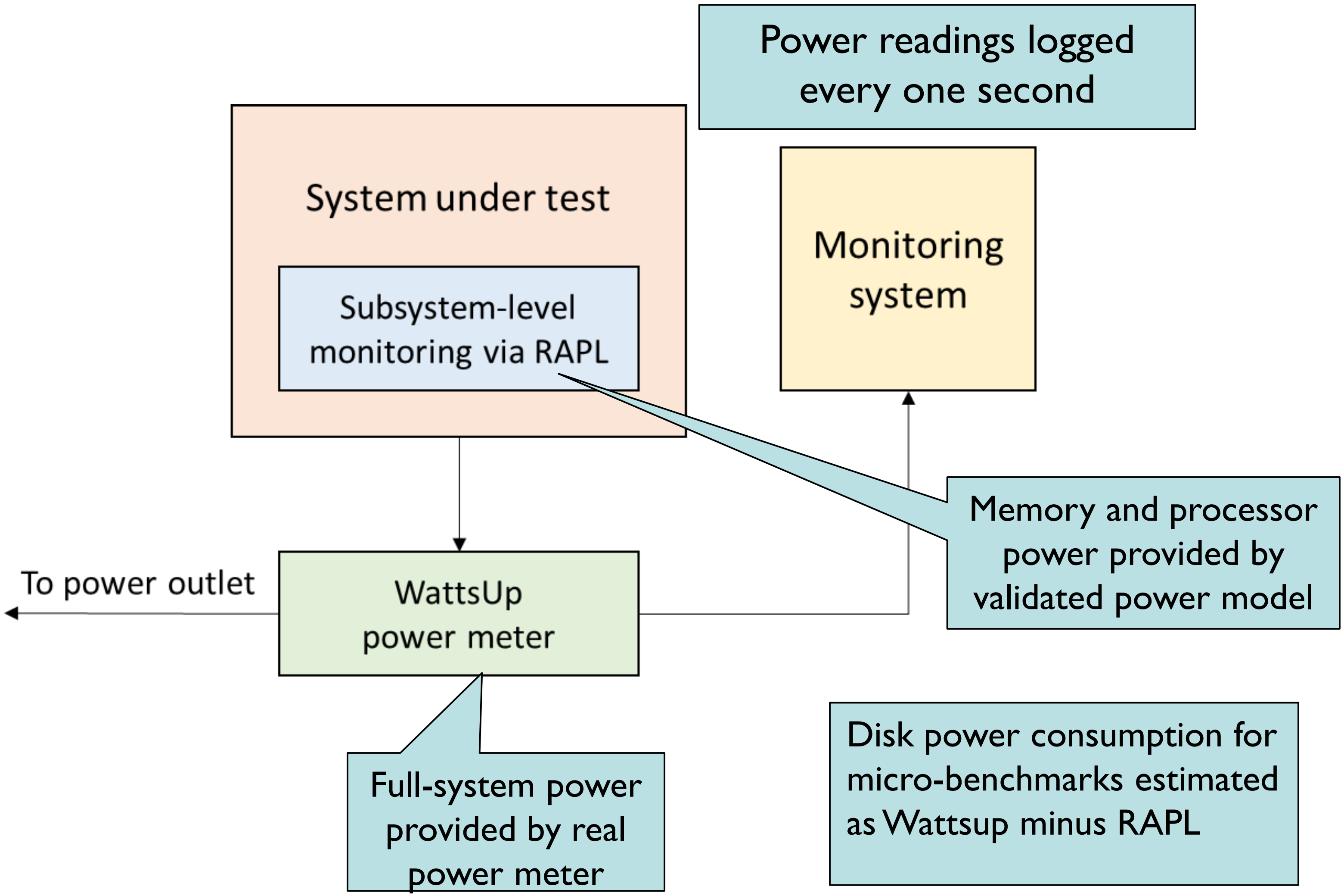
APPLICATION



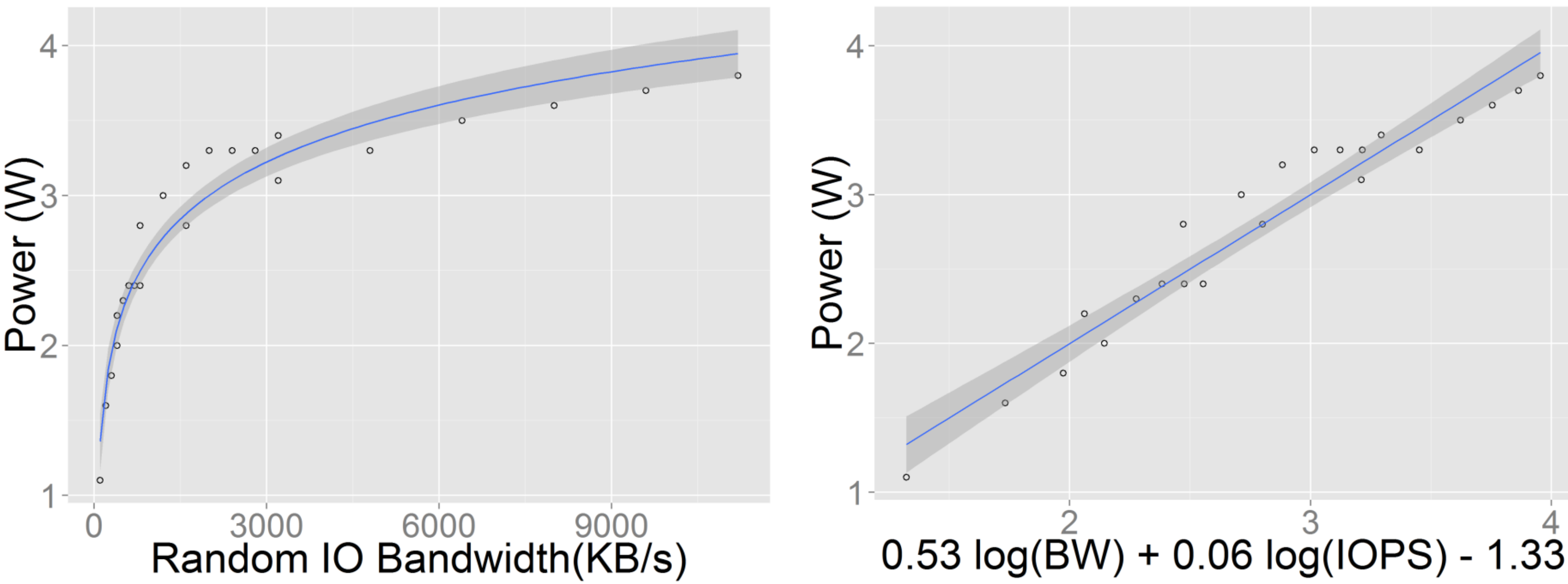
MPAS Ocean Simulation

End goal: Identify eddies in the ocean

DATA COLLECTION

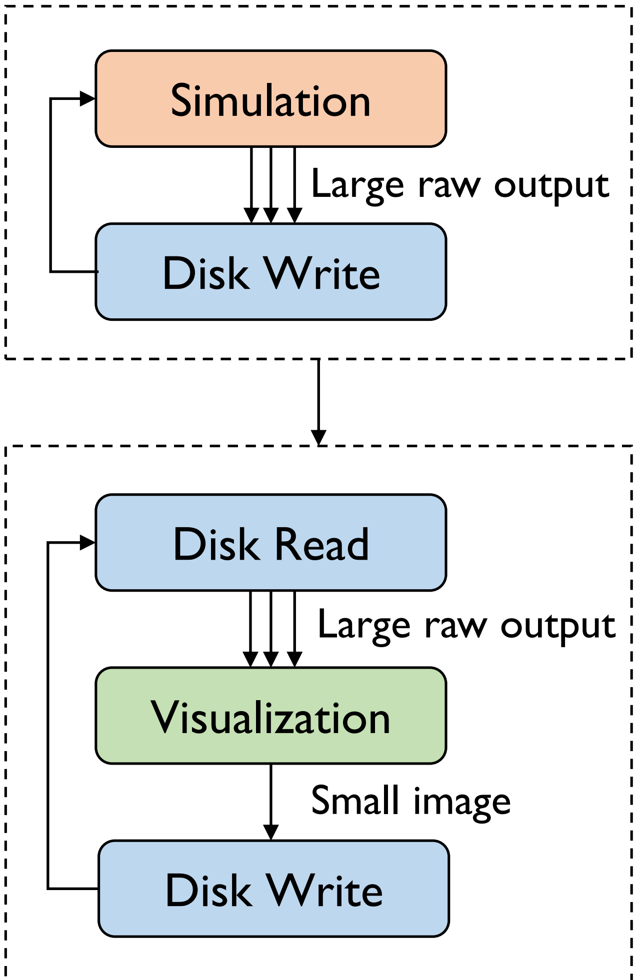


DISK POWER MODEL

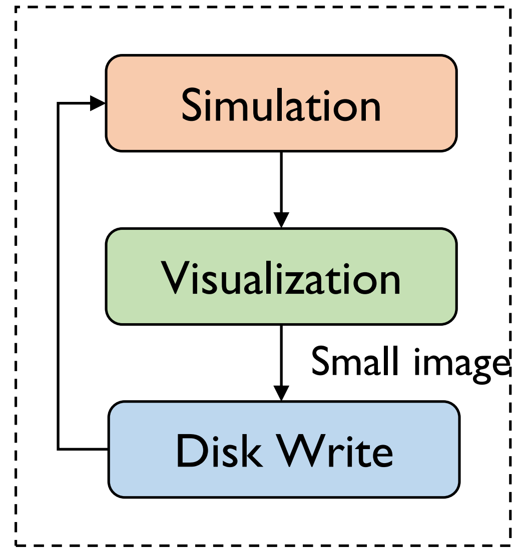


- I/O statistics collected from *iostat*
- Number of I/O operations and the amount of data written affects power consumption of the disk

Pipelines Studied



Post-processing



In-situ

Traditional Post-Processing: Post-processing without any sampling
Modern Post-Processing: Post-processing with temporal sampling (write output every few iterations – here every 24 itereations)
In-situ: Produce images on the fly *and* do so only every few iterations

Hardware Platform

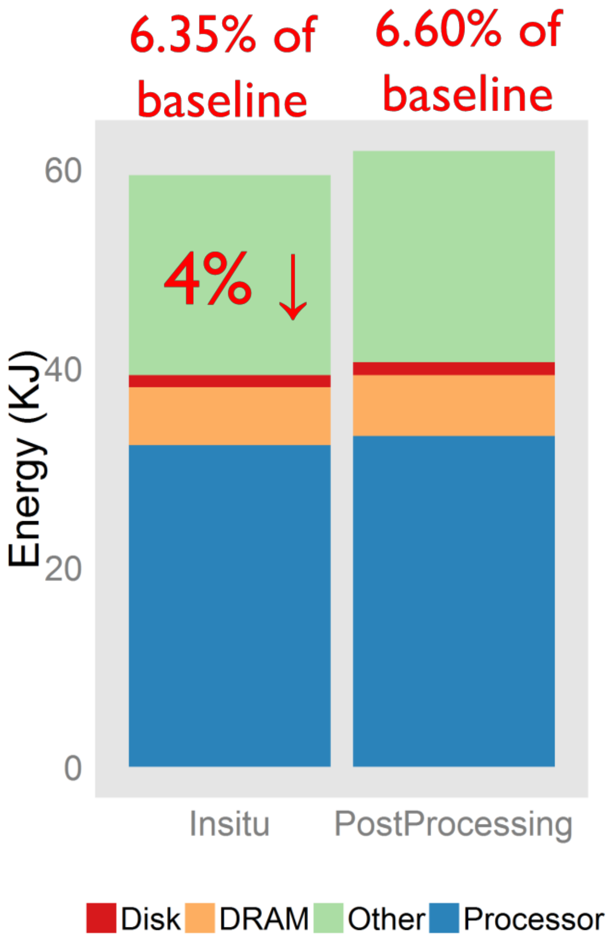
CPU	2x Intel Xeon E5-2665
CPU frequency	2.4 GHz
Last-level cache	20 MB
Memory	4x 16GB DDR3-1333
Memory size	64 GB
Hard disk	Seagate 7200rpm disk
Storage size	500GB
Disk bandwidth	6.0 Gbps

Hardware configuration

OUTLINE

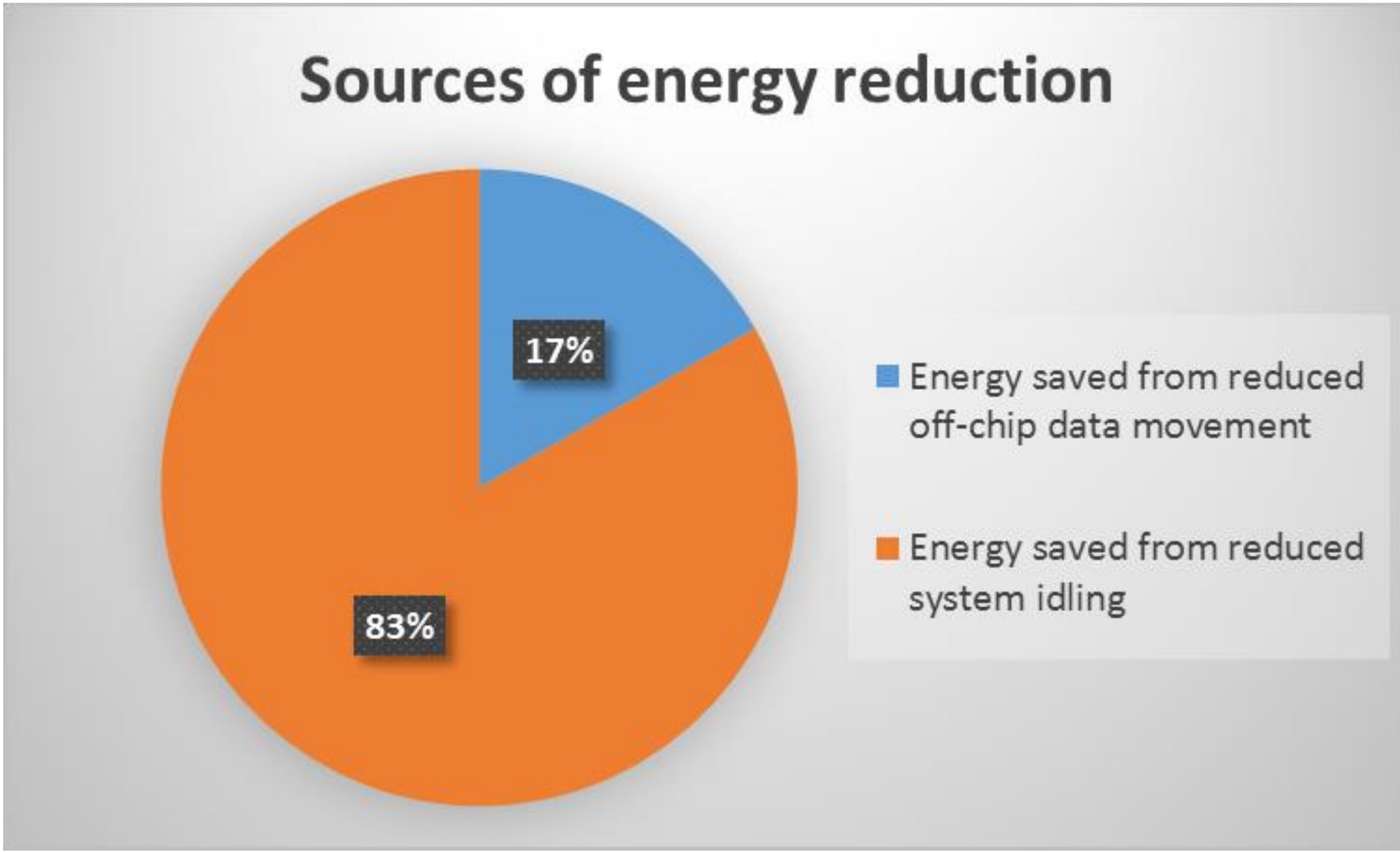
- Introduction
- Objective
- Methodology
- Results
- Conclusion and Future Work

Results: Single-Node Energy Comparison

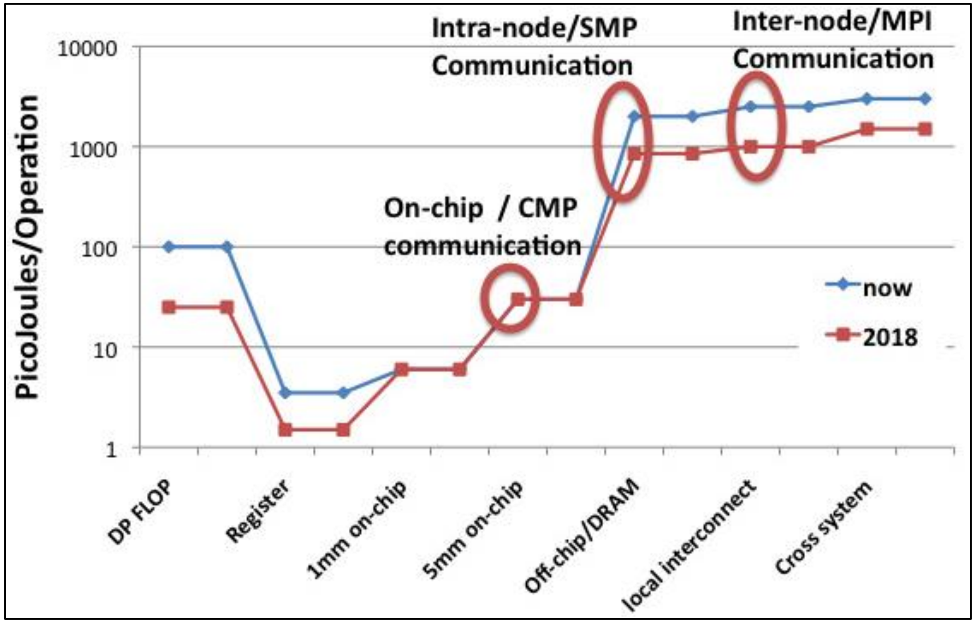


- In-situ consumes 4% less energy than *modern* post-processing
- Compared to *traditional* post-processing, both pipelines consume 93% lower energy
  - Traditional post-processing no longer accepted as baseline
    - Storage limitations won't let it happen
  - True baseline somewhere between traditional and modern post-processing

Results: Single-Node Energy Comparison



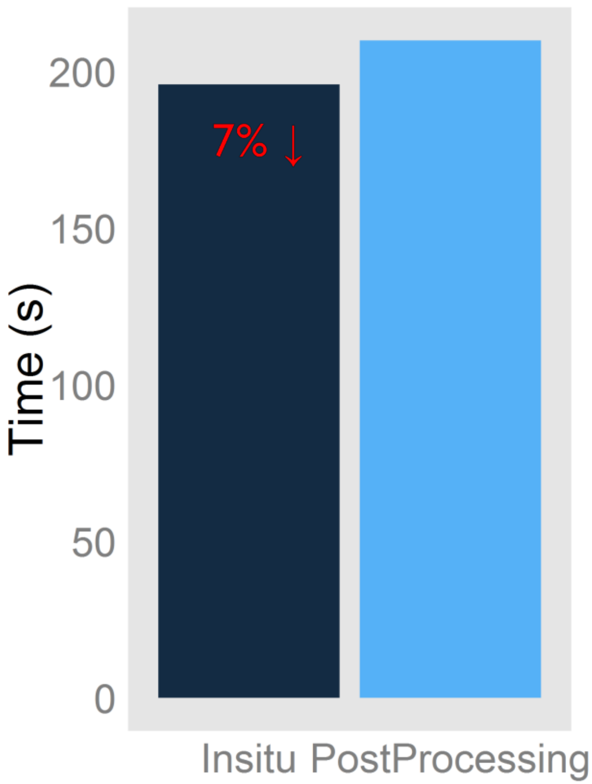
POWER/ENERGY CHALLENGE



- Energy consumed for moving a bit increases as we move down the memory hierarchy
  - Off-chip transfers costs nearly 100 times as much energy as on-chip transfers

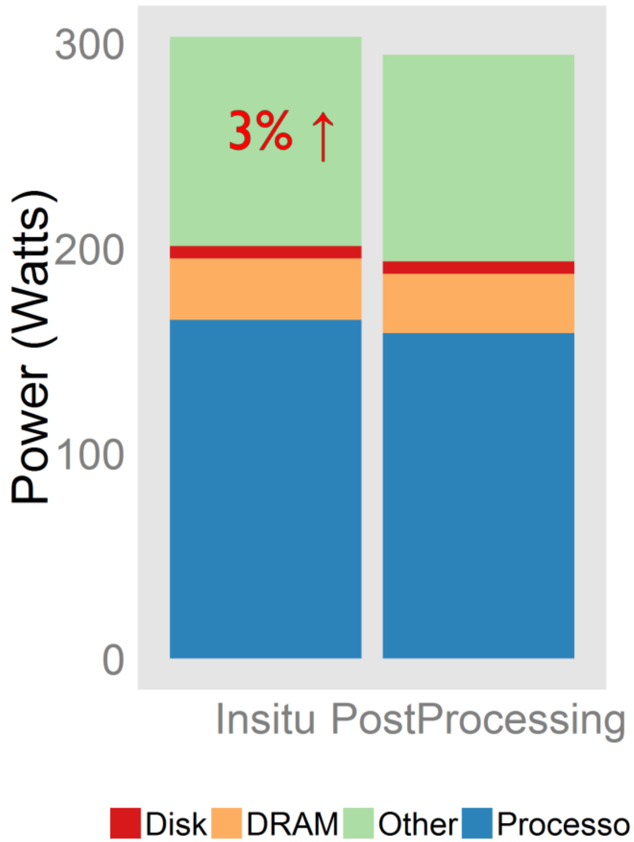
Image source: J.Shalf et al., "Exascale Computing Technology Challenges", VECPAR 2010

Results: Single-Node Performance Comparison



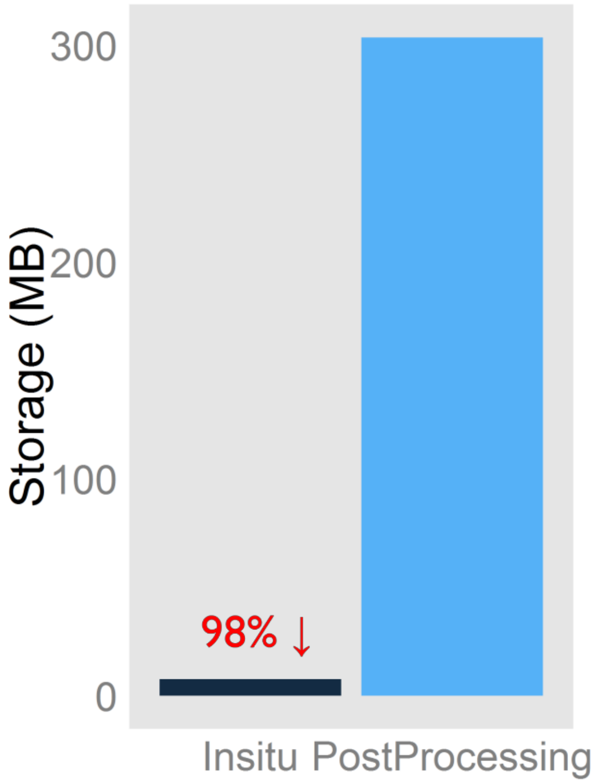
- In-situ consumes 7% lower execution time than modern post-processing
  - Reduced I/O wait time
- Difference will be significant for a HPC system
  - Details later

RESULTS: SINGLE-NODE POWER COMPARISON



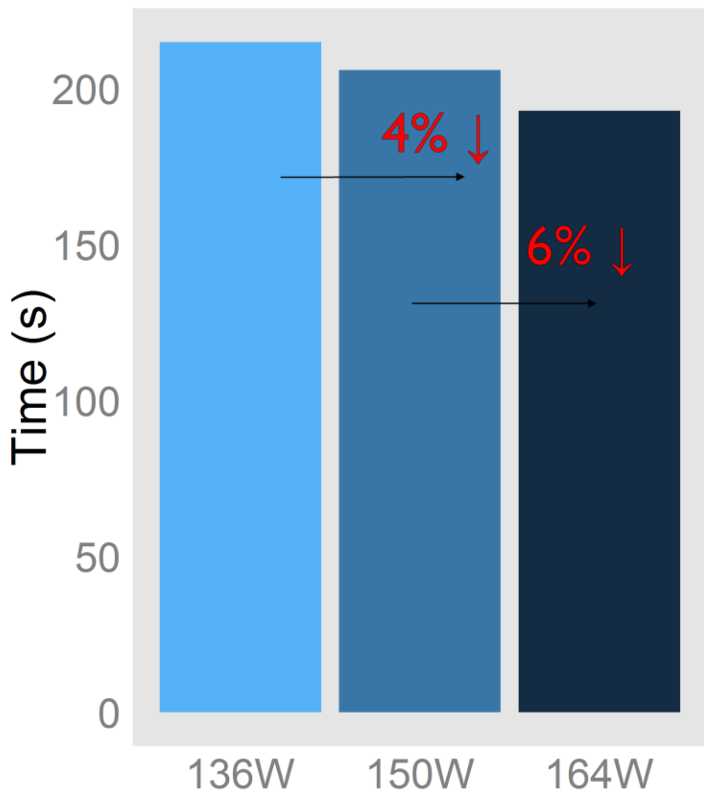
- In-situ consumes 3% more power than *modern* post-processing
  - Difficult trade-off choice
- Might not be the same for a supercomputer
  - Details later

RESULTS: SINGLE-NODE STORAGE REQUIREMENTS



- ~97.5% lower storage requirement for the in-situ pipeline
    - Implies smaller storage cluster
    - Implies lower power consumption

Redistributing Storage Power to Compute Nodes



- Assuming reduced storage nodes results in 10% of total power redirected to computer nodes
  - Performance improves by up to 6% for MPAS-O
    - Data from power capping experiments with RAPL

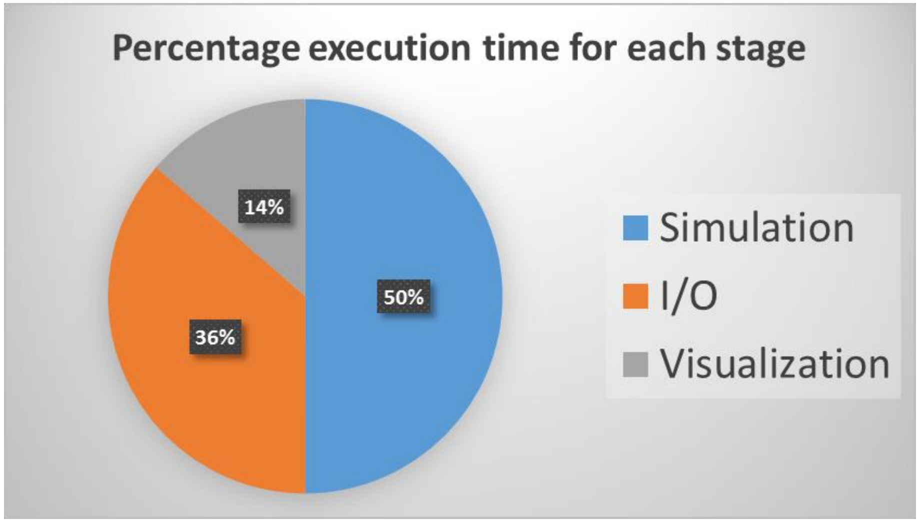
EXPECTATIONS FOR A SUPERCOMPUTER

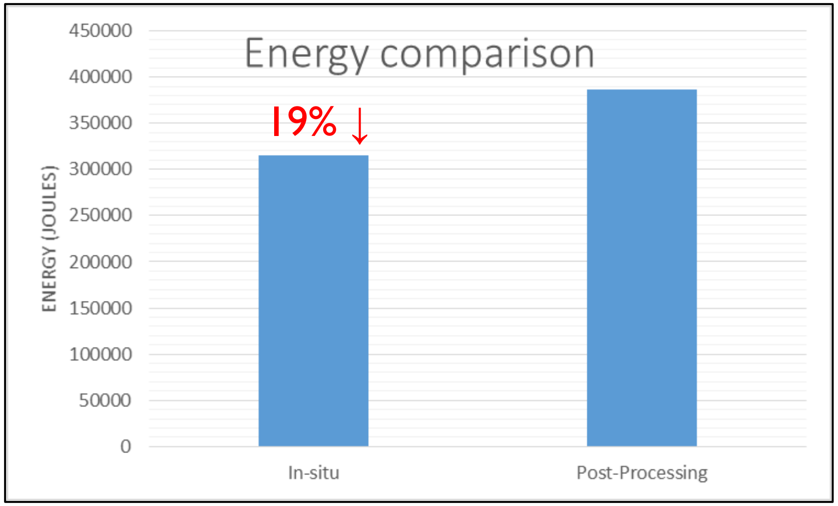
- Increased I/O wait time
  - Storage separated from compute by network
  - Longer execution time and corresponding increase in energy
- Additional energy consumption from data movement through the network
  - No data transfer via network cables in single-node
- Power/energy overhead for storage higher
  - Separate cluster for storage → additional CPUs, memory, cooling etc.
  - Storage sub-system shared with compute sub-system in single-node

Results at Scale: Hardware Platform

- Caddy supercomputer with a dedicated Lustre file system used for profiling
- Compute nodes
  - 64 nodes out of 150 nodes used in these experiments
    - Each node contains 2x Intel Xeon E5-2670 and 64 GB of RAM
  - Nominal power consumption
    - 6000 W (idle) to 20000 W (workload such as MPAS)
- Storage nodes
  - 5 nodes configured as 1 master + 2 MDS + 2 OSS
  - 1 RAID storage per MDS and OSS
  - Nominal power consumption
    - 2500W (idle) to 2800W (active)

RESULTS AT SCALE: ENERGY COMPARISON





Real measurements on Caddy

Partial measurement on Caddy and extrapolation from spec sheets

OUTLINE

- Introduction
- Objective
- Methodology
- Results
- Conclusion and Future Work

FINDINGS

- Most energy savings come from reducing system idling (i.e., from reduced I/O wait time)
- Further savings possible if we can reduced size of the storage nodes

Conclusion

- In-situ visualization offers the following advantages:
  - Reduced energy consumption (by reducing system idling or I/O wait time)
  - Reduced power (by using fewer storage nodes)
  - Improved performance (by reducing I/O wait time and by making more power available for compute nodes)

Future Directions

- Enhancing HPC systems
  - Flash buffers and SSDs can reduce I/O wait time
    - Downside: Introducing more components can increase power consumption
- HPC system design changes
  - Bringing storage nodes and compute nodes together
    - Similar to Memory in Processor or Processor in Memory concepts in the computer architecture community
- Runtime system changes
  - Energy proportional computing and storage
    - Putting compute nodes to sleep states during I/O
    - Putting some storage nodes to deep sleep state when bandwidth and storage requirements are lower